

# Methodischer Bias retrospektiver Follow-up-Untersuchungen und sein Einfluss auf die Erhebung von Qualitätssicherungsdaten

## Methodical Bias of Retrospective Follow-up Investigations and Impact on the Collection of Quality Assurance Data

### Autoren

T. Karn, U. Holtrich, E. Ruckhäberle, L. Hanker, A. Schlieter, C. Solbach, R. Gätje, M. Kaufmann, A. Rody

### Institut

Klinik für Gynäkologie und Geburtshilfe, J.W.Goethe-Universität, Frankfurt

### Schlüsselwörter

- retrospektive Kohortenstudien
- Information-Bias
- Selection-Bias

### Key words

- retrospective cohort studies
- information bias
- selection bias
- loss of follow-up

### Zusammenfassung

**Einführung:** Derzeit kann nicht prospektiv belegt werden, dass die Einhaltung operativer Qualitätsparameter im Rahmen der Qualitätssicherung des Mammakarzinoms mit einer Verbesserung des krankheitsfreien Überlebens und Gesamtüberlebens verbunden ist. Aufgrund eines fehlenden nationalen klinischen Krebsregisters können Daten zum krankheitsfreien Überleben bzw. Gesamtüberleben vielfach nur durch retrospektive Erhebungen erfasst werden. Allerdings muss aufgrund von Selektions- und Observationsfehlern angenommen werden, dass die entsprechenden Daten eine hohe Fehlerrate aufweisen, deren Ausmaß schwierig abzuschätzen ist.

**Material und Methode:** Von n = 710 Mammakarzinompatientinnen, die im Zeitraum von 1997–1999 an der Universitätsfrauenklinik Frankfurt behandelt wurden, wurden Überlebensdaten retrospektiv erfasst. Eine vollständige Datenerfassung konnte nur für 471 Patientinnen erzielt werden. Der Einfluss von Selektions- und Information-Bias auf die Datenqualität und das Überleben der Patientinnen sollte überprüft werden.

**Ergebnisse:** Das Untersuchungskollektiv zeichnete sich im Vergleich zu großen Populationskollektiven durch einen höheren Anteil an Risikopatientinnen aus (nodal positiv, höhere Tumorgroße). Überraschenderweise war die mediane Beobachtungszeit in der Gruppe der verstorbenen Patientinnen 1,7-fach größer als bei den zensierten Fällen, während sich in Vergleichsuntersuchungen ein umgekehrter Zusammenhang zeigt. Dies lässt auf einen „*Information-Bias*“ schließen, bei dem v.a. bei krankheitsfreien Patientinnen ein „loss of follow-up“ auftrat. Die Einflussgröße dieses Bias wurde daher durch eine Simulation abgeschätzt. Als Ergebnis wurde im Vergleich zur retrospektiven Studie ein möglicher Unterschied bez. des 5-Jahres-Überlebens von 13,9% ( $73,5 \pm 3,1\%$  vs.  $87,4 \pm 1,5\%$ ) erhalten.

### Abstract

**Background:** It seems reasonable to assume that standardized quality assurance and adherence to guidelines in the treatment of breast cancer will lead to an improvement in disease-free and overall survival. However there are still no prospective studies demonstrating this relationship. Because of the absence of a national clinical cancer registry in Germany, survival data are often collected using retrospective studies which are prone to different errors and biases.

**Material and Methods:** Survival information was retrospectively collected on 710 breast cancer patients undergoing surgical treatment at the University of Frankfurt in the years 1997–1999. Complete data were obtained for 471 patients. The influence of selection bias and information bias on data quality was analyzed.

**Results:** The study cohort was characterized by an elevated rate of patients with higher risk (lymph node positive and larger tumor size) compared to large population-based reference cohorts. Interestingly, the median follow-up time for the group of deceased patients was 1.7 times longer than that of censored cases. This is in contrast to other studies demonstrating an inverse relationship, whereby the proportion of drop-outs from follow-up is higher among disease-free patients in the study, and suggests an information bias. A simulation experiment was performed to estimate the influence of this bias and a possible difference of 13.9% in the 5-year overall survival was suggested ( $73.5 \pm 3.1\%$  vs.  $87.4 \pm 1.5\%$ ).

**Conclusion:** Retrospective studies on survival are influenced by several confounding variables in addition to classical clinical parameters. These factors complicate comparisons with reference cohorts and could result in a misleading assumption of a worse quality of treatment.

**eingereicht** 30. 10. 2009

**akzeptiert** 17. 11. 2009

### Bibliografie

**DOI** <http://dx.doi.org/10.1055/s-0029-1240728>  
 Geburtsh Frauenheilk 2010; 70: 41–46 © Georg Thieme  
 Verlag KG Stuttgart · New York ·  
 ISSN 0016-5751

### Korrespondenzadresse

**Priv.-Doz. Dr. med. Achim Rody**  
 Klinik für Gynäkologie  
 und Geburtshilfe  
 Goethe Universität  
 Theodor-Stern-Kai 7  
 60590 Frankfurt  
 achim.rody@  
 em.uni-frankfurt.de

**Zusammenfassung:** Die retrospektive Überlebensanalyse von Patientinnen unterliegt zahlreichen Confoundern, die unabhängig von den klassischen klinischen Parametern, eine Vergleichbarkeit mit anderen, großen Kollektiven erschwert und somit unter Umständen schlechtere Therapiequalität widerspiegelt, die jedoch nicht zwangsläufig real anzutreffen ist.

## Einleitung

Qualitätssicherung im Rahmen der Versorgung von Brustkrebspatienten ist ein wichtiges Instrument zur Verbesserung diagnostischer, therapeutischer und strukturell-organisatorischer Maßnahmen. In diesem Zusammenhang werden im Rahmen der Offenlegung der Qualitätssicherungsdaten sowie von Disease-Management- und Zertifizierungsprogrammen v.a. primäre therapeutische Qualitätssicherungsdaten wie z.B. die Rate brusterhaltender Therapien, die Angabe des Resektionsrands und die Anzahl entfernter Lymphknoten erfasst und hinsichtlich aktueller Leitlinien [1] ausgewertet. Das National Institute for Health and Clinical Excellence (NICE) hat in seinen Guidelines allen Brustzentren zudem empfohlen, die zentrumsspezifische Rate an Lokal- und Axillarezidiven zu erfassen [2].

Die entscheidenden Surrogatmarker für den Erfolg aller Bemühungen sind jedoch eine Verbesserung des krankheitsfreien Überlebens sowie des Gesamtüberlebens. Allerdings liegen bis dato keine prospektiv erhobenen Daten vor, die einen Zusammenhang von Qualitätssicherungsparametern und einer entsprechenden Verbesserung des Patienten-Outcomes zeigen, auch wenn dieser Zusammenhang als plausibel angenommen werden muss. Zahlreiche Brustzentren versuchen, diesen Zusammenhang durch retrospektive Datenerhebungen darzustellen. Solche Untersuchungen können jedoch aufgrund ihrer immanenten Struktur zu verschiedenen Arten von Bias führen.

Bei einer solchen Datenerhebung handelt es sich grundsätzlich um eine retrospektive Kohortenstudie. Diese Studienform lässt sich methodisch folgendermaßen charakterisieren:

- ▶ In einer Kohortenstudie wird eine Gesamtkohorte (hier: Patientinnen mit operiertem primärem Mammakarzinom) hinsichtlich des Auftretens eines Ereignisses (hier Rezidiv) verfolgt [3].
- ▶ Die Vorteile von Kohortenstudien gegenüber Case-Control-Studien liegen darin, dass die Inzidenz direkt ermittelt wird und z.B. kein Recall-Bias durch unterschiedliche Erinnerung auftreten kann.
- ▶ Ein weiterer Vorteil retrospektiver Kohortenstudien liegt darin, dass die Zeit bis zum Auftreten von bestimmten Ereignissen nicht abgewartet werden muss. Nachteile sind andererseits, dass die erhobenen Daten keine Gewähr auf Verlässlichkeit bieten [4].

Im Gegensatz zu prospektiv randomisierten Studien sind jedoch im Falle einer retrospektiven Kohortenstudie 2 wesentliche kritische Einflüsse zu befürchten [5]:

1. **Selection-Bias:** Die Möglichkeit, dass „Cases“, also Fälle mit Ereignis, mit einer höheren Wahrscheinlichkeit in die Studie aufgenommen werden, als sie in der Gesamtgruppe der Erkrankten vorliegen.
2. **Information- oder Observation-Bias:** Dies bedeutet, dass ein Verlust von Informationen, z.B. durch „loss of follow-up“, häufiger in einer der Gruppen (Cases vs. Kontrollen) auftritt.

Ein Selection-Bias kann in der Praxis z.B. dadurch auftreten, dass eine universitäre Einrichtung, die besonders aktuelle Therapie-

konzepte anbietet, einen größeren Anteil an Patienten mit einer *per se* schlechteren Prognose aufweisen könnte. Solch ein Bias ließe sich jedoch durch eine Auswertung der klinischen Daten der Patientinnen überprüfen [6]. Kritischer sind dagegen verschiedene Typen von Information-Bias wie z.B. der „surveillance bias“: Oft kommen gerade die Patientinnen mit einem höheren Rezidivrisiko häufiger zu regelmäßigen Kontrolluntersuchungen in die Klinik, in der sie primär behandelt wurden. Auch dieser Bias sollte sich überprüfen lassen. In diesem Fall sollten Patientinnen mit einem aufgrund ihrer klinischen Daten höheren Rezidivrisiko durchschnittlich längere Follow-up-Zeiten aufweisen als Low-Risk-Patientinnen. Ein sehr kritischer Information-Bias kann dadurch entstehen, dass ein „loss of follow-up“ häufiger in der Gruppe ohne Ereignis auftritt [7]. So könnten im Extremfall alle vollständig geheilten Patientinnen einen frühen „dropout“ aufweisen und z.B. schon nach den ersten frühen Nachuntersuchungen aufgrund fehlender weiterer Informationen zensiert werden. Im Follow-up würden anschließend nur noch jene Patientinnen erscheinen, die früher oder später ein Rezidiv aufweisen. Ein solcher Effekt würde eine äußerst schlechte Prognose des Patientenkollektivs der Studie vorspiegeln. Eine kürzere durchschnittliche Beobachtungszeit der zensierten Patientinnen im Vergleich zu den verstorbenen Patientinnen oder jenen mit einem Ereignis könnte zwar einen Hinweis auf einen solchen Bias liefern. Die Stärke dieses äußerst kritischen Effekts lässt sich jedoch nicht aus den Daten ableiten.

Um zu untersuchen, inwieweit sich die genannten potenziellen Fehlerquellen in der Praxis störend auf das Ergebnis einer retrospektiven Studie auswirken, wurde eine empirische Analyse des Follow-ups von Patientinnen durchgeführt, die in den Jahren 1997–1999 an der Universitätsfrauenklinik behandelt wurden.

## Patienten und Methoden

▶ Patientinnen mit einem primären Mammakarzinom, die an der Universitätsfrauenklinik Frankfurt in den Jahren 1997–1999 behandelt wurden und für die ein vollständiger primärer Datensatz (bestehend aus Operationstyp, TNM-Stadium, Hormonrezeptorstatus und vollständige Informationen zur abgeschlossenen adjuvanten Therapie, vorhandene Kontaktdaten) vorlag, wurden in die retrospektive Follow-up-Erhebung aufgenommen. Eine vollständige Datenerfassung von insgesamt 710 in diesem Zeitraum behandelten Patientinnen lag nur für 471 Frauen vor. Mithilfe eines strukturierten Fragebogens, der an die betreuenden Ärzte verschickt wurde, konnten entsprechende Nachbeobachtungsdaten der Patientinnen, mit dem Ziel, einen Follow-up-Zeitraum von mindestens 5 Jahren zu erfassen, abgefragt werden. Die Antwortrate betrug 48,6% (n = 229). Die erhaltenen Daten wurden mit jenen von Populationsstudien in Finnland [8] und den USA [9] verglichen, um einen möglichen Selection-Bias aufzudecken. Weiterhin wurden Simulationsexperimente zum Einfluss eines Information-Bias durchgeführt. Die Untersuchungen soll-

**Tab. 1** Vergleich der klinischen Parameter des Frankfurter Kollektivs mit Populationsstudien aus Finnland und den USA.

		FinProg (1991–1992)		SEER (1991–1992)		Frankfurt (1997–1999)		p-Value ( $\chi^2$ -Test) Frankfurt vs.	
		n	%	n	%	n	%	FinProg	SEER
Alter	< 50	567	27,8	6984	27,2	192	40,8		
	≥ 50	1469	72,2	18738	72,8	279	59,2	< 0,001	< 0,001
Tumorgröße	< 2 cm (T1)	1208	62,1	15001	63,2	240	52,6		
	≥ 2 cm	737	37,9	8743	36,8	216	47,4	< 0,001	< 0,001
Nodalstatus	LNN	1239	65,8	14706	66,3	242	54,1		
	N1	643	34,2	7476	33,7	205	45,9	< 0,001	< 0,001
histologisches Grading	G1	394	26	2068	12,3	70	16		
	G2	716	47,3	6991	41,7	232	53		
	G3	403	26,6	7713	46	136	31,1	< 0,001	< 0,001
Östrogenrezeptor-Status	positiv	1082	69,9	15359	76,1	309	72,4		
	negativ	466	30,1	4835	23,9	118	27,6	n. s.	n. s.
Progesteronrezeptor-Status	positiv	940	60,8	12955	65,8	261	61,6		
	negativ	607	39,2	6729	34,2	163	38,4	n. s.	n. s.

ten 3 Fragen überprüfen, die sich aufgrund der oben genannten potenziellen Fehlerquellen ergeben:

1. Besteht ein Selection-Bias bezüglich der Patientencharakteristika zwischen dem Kollektiv der Universitätsfrauenklinik Frankfurt und Vergleichskollektiven? Um einen potenziellen Selection-Bias zu identifizieren, wurden die Kollektive hinsichtlich der Verteilung ihrer klinischen Parameter verglichen und mögliche signifikante Unterschiede durch  $\chi^2$ -Test untersucht. Verwendet wurden hierfür Vergleichskollektive von 2 Populationsstudien aus Finnland (FinProg) [8] und den USA (SEER) [9].
2. Besteht ein Information-Bias im Follow-up der Patientinnen der Universitätsfrauenklinik Frankfurt hinsichtlich des Rezidivrisikos? Zur Detektion eines potenziellen Information-Bias hinsichtlich des Rezidivrisikos wurden die Beobachtungszeiten von Patientinnen, die im Laufe der Nachbeobachtung verstorben waren, mit den Beobachtungszeiten zensierter Patientinnen verglichen.
3. Wie groß ist der mögliche Einfluss eines potenziellen Information-Bias auf das Ergebnis der Follow-up-Analysen? Diese Frage sollte mithilfe einer Simulation geklärt werden. Es wurde hierzu hypothetisch postuliert, dass sämtliche zensierten Patientinnen bei Abschluss der Studie noch am Leben wären. Der Einfluss dieser Annahme auf die Rate des Gesamtüberlebens wurde überprüft.

## Ergebnisse



### Vergleich der Zusammensetzung des Studienkollektivs mit Referenzkollektiven

In **Tab. 1** wird der Vergleich der klinischen Parameter des Frankfurter Kollektivs mit 2 Populationsstudien aus Finnland [8] und den USA [9] dargestellt. Das Frankfurter Kollektiv zeichnet sich durch einen größeren Anteil jüngerer Patientinnen (Alter < 50 Jahre: 40,8 vs. 27,2% bzw. 27,8%), solchen mit größeren Tumoren (47 vs. 37% bzw. 38% über 2 cm) sowie positivem Nodalstatus (46 vs. 34%) aus. All diese Parameter unterscheiden sich hoch signifikant zu den Vergleichskollektiven im Gegensatz zum Hormonrezeptor-Status, wo weder bez. Östrogen- noch Progesteronrezeptor ein signifikanter Unterschied beobachtet wurde.

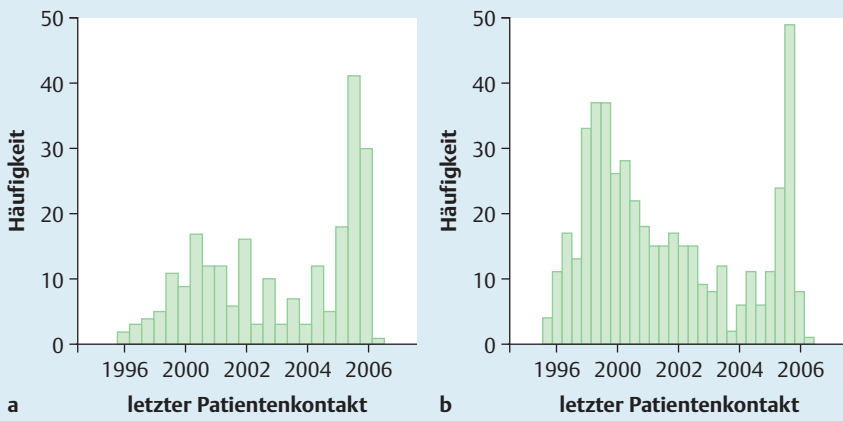
Während für das finnische Kollektiv bez. des histologischen Gradings ein geringerer Anteil an G3-Tumoren und ein höherer Anteil an G1-Tumoren als beim Frankfurter Kollektiv bestimmt wurde, zeigt der Vergleich mit dem US-Kollektiv ein umgekehrtes Ergebnis.

Insgesamt lässt sich aufgrund des jüngeren Alters der Patientinnen sowie des höheren Anteils an nodal-positiven und größeren Tumoren eine schlechtere Prognose des Frankfurter Kollektivs vermuten.

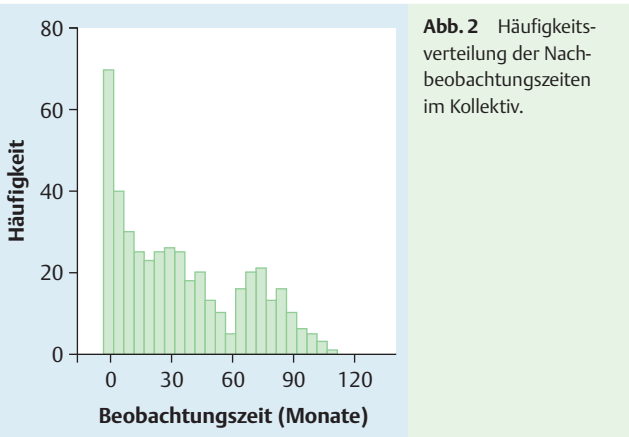
### Analyse des Patientenkollektivs der Studie hinsichtlich eines potenziellen Information-Bias in den Follow-up-Daten

☛ **Abb. 1 a** zeigt die Verteilung der Angaben zum Datum des letzten Patientenkontakts, das von den niedergelassenen Arztpraxen zurückgemeldet wurde. In **Abb. 1 b** wurden zusätzlich auch jene Patientinnen eingeschlossen, von denen nur aus der Universitätsklinik Frankfurt stammende Informationen vorlagen. Aufgrund dieser Daten und dem jeweils zugehörigen Erstdiagnosedatum lässt sich die in **Abb. 2** gezeigte Verteilung der Beobachtungszeiten der Patientinnen ermitteln. Neben einer großen Zahl von Fällen, für die nur sehr knappe Follow-up-Informationen (< 6 Monate) vorliegen, zeigt sich hier eine Häufung von Fällen mit einer Beobachtungszeit von 2–3 Jahren sowie ein 2., deutlicher Peak bei etwa 6 Jahren, der sich mit dem Abstand des Diagnosezeitraums (1997–1999) mit der im Rahmen dieser Arbeit durchgeführten Befragung der niedergelassenen Frauenärzte (2005) deckt (s. a. **Abb. 1 a**).

In **Abb. 3 a** wird die Gesamtbeobachtungszeit bei Patientinnen verglichen, die entweder verstarben oder zum Ende der Beobachtungszeit noch am Leben waren und zensiert wurden. Die mediane Beobachtungszeit in der Gruppe der verstorbenen Patientinnen ist mit 40 Monaten sehr viel größer (1,7-fach) als jene der zensierten Patientinnen mit 24 Monaten. Ein in **Abb. 3 b** dargestelltes unabhängiges Vergleichskollektiv aus einer Metaanalyse von Kohorten mit Genchipdaten [10] zeigt dagegen ein umgekehrtes Bild. Dort ist die mediane Beobachtungszeit in der Gruppe der zensierten Patientinnen (123 Monate) mehr als doppelt so groß wie jene der verstorbenen Patientinnen (56 Monate). Auch bei Einzelanalysen der hier gemeinsam dargestellten 3 Kohorten [11–13] zeigte sich durchgängig eine 2- bis 3-fach längere



**Abb. 1 a und b** Häufigkeitsverteilung der Kalenderdaten des letzten Patientenkontakts. In **a** wurden nur die Informationen eingeschlossen, die von niedergelassenen Arztpraxen zurückgemeldet wurden, wohingegen in **b** auch jene Patientinnen eingeschlossen wurden, von denen nur aus der Universitätsklinik Frankfurt stammende Informationen vorlagen.



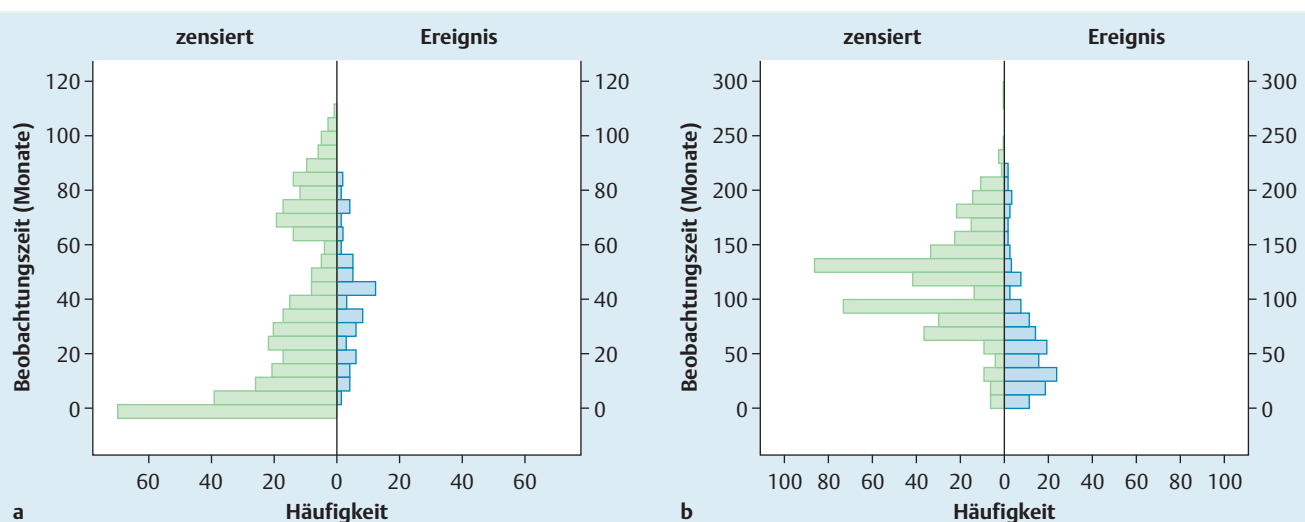
**Abb. 2** Häufigkeitsverteilung der Nachbeobachtungszeiten im Kollektiv.

mediane Überlebenszeit bei den zensierten Patientinnen (nicht gezeigt). Dieser Vergleich deutet darauf hin, dass in unserer Untersuchung ein Information-Bias vorliegt, bei dem ein „loss of fol-

low-up“ häufiger in der Gruppe der ohne Ereignis zensierten Patientinnen auftrat.

### Simulationsexperiment zur Überprüfung des Einflusses eines möglichen Information-Bias

Als nächstes wurde ein Simulationsexperiment durchgeführt, um die Größe des vermuteten Information-Bias im untersuchten Kollektiv zu bestimmen. Die Erstdiagnose und Operation der Patientinnen des Kollektivs erfolgte in den Jahren 1997–1999, während die letzten Follow-up-Informationen im Februar 2006 erhalten wurden. Daraus ergibt sich als theoretisch mögliches maximales Follow-up der Patientinnen ein Zeitraum zwischen 74 und 109 Monaten. Aus diesem Grund wurde sich in der folgenden Simulation auf ein Follow-up von 75 Monaten beschränkt und sämtliche Fälle ohne Ereignis wurden zu diesem Zeitpunkt zensiert. Verwendet man die aus der retrospektiven Studie vorliegenden Follow-up-Informationen für eine Kaplan-Meier-Analyse, so ergibt sich für das Gesamtüberleben (OAS, Overall Survival) der Patientinnen das in der unteren Kurve („A“) in **Abb. 4** gezeigte Bild. Die 5-Jahres-Überlebensrate würde in diesem Fall



**Abb. 3 a und b** Vergleich der Nachbeobachtungszeiten von Patientinnen mit Ereignis und zensierten Fällen. **a** zeigt die Daten des hier untersuchten Kollektivs mit einer längeren medianen Beobachtungszeit bei Patientinnen mit Ereignis (40 Monate) als bei zensierten Fällen (24 Monate). In **b** sind Da-

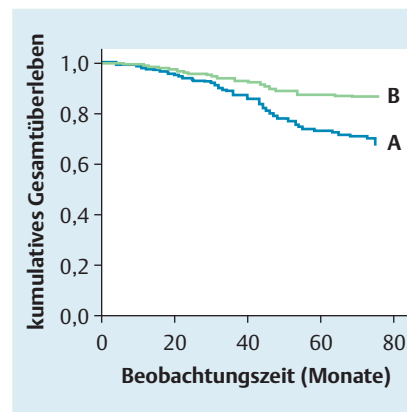
ten eines Vergleichskollektivs gezeigt mit einer kürzeren medianen Beobachtungszeit in der Gruppe mit Ereignis (56 Monate) als bei zensierten Fällen (123 Monate).

bei  $73,5 \pm 3,1\%$  liegen (Abb. 4). Postuliert man jedoch hypothetisch, dass sämtliche Patientinnen, die vor diesem Zeitpunkt aufgrund fehlender Informationen aus dem Follow-up zensiert wurden, auch nach 75 Monaten noch am Leben sind, so ergäbe sich der in der oberen Kurve („B“) in Abb. 4 gezeigte Verlauf. Unter diesen Umständen läge die 5-Jahres-Überlebensrate deutlich höher bei  $87,4 \pm 1,5\%$ . Da natürlich davon auszugehen ist, dass nicht alle zensierten Patientinnen ein Überleben von 75 Monaten aufweisen, geben die beiden dargestellten Ergebnisse mit einem Unterschied von 13,9% somit die Extreme wieder, zwischen denen das wahre Ergebnis einzuordnen sein sollte. Die 5-Jahres-Überlebensraten der Vergleichskollektive aus Finnland [8] und den USA [9] betragen  $78,7 \pm 1,8\%$  und  $79,7 \pm 0,5\%$  und liegen somit 5–6% über dem direkt erhaltenen Wert und 8–9% unter dem der hypothetischen Simulation. Schließt man in die Überlegungen mit ein, dass wie oben gezeigt die klinischen Parameter des Frankfurter Kollektivs eine schlechtere Prognose vorhersagen, wäre ein Wert zwischen 74 und 78% für das 5-Jahres-Überleben dieses Kollektivs denkbar und der Information-Bias würde demnach zu einem Fehler unter 5% führen.

Analoge Ergebnisse erhält man auch nach Stratifikation der Proben aufgrund von klinischen Parametern. So zeigt die Abb. 5 beispielhaft die Kaplan-Meier-Überlebensanalyse bei Stratifikation der Patientinnen nach Tumorstadium T1–T4. Hier ist in Abb. 5a das Ergebnis der vorhandenen Follow-up-Informationen der Studie dargestellt und in Abb. 5b das simulierte Ergebnis, bei hypothetisch postuliertem Überleben aller zensierten Patientinnen. Auch hier ist zu vermuten, dass die Wahrheit zwischen diesen beiden Extremen liegt.

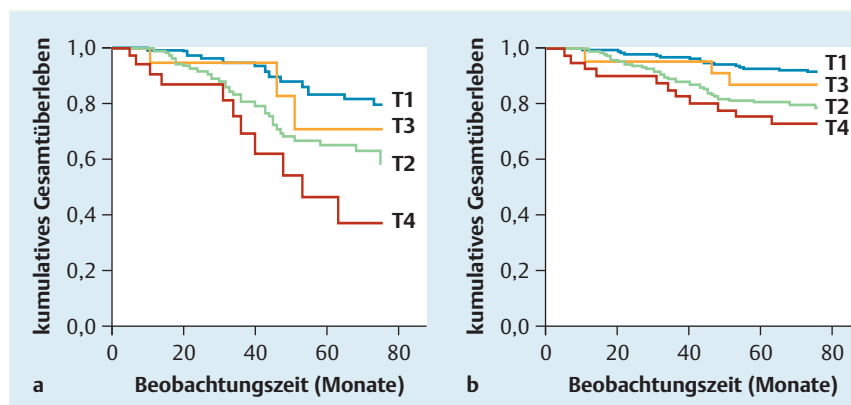
## Diskussion

Die Beurteilung der Therapiequalität bei Patientinnen mit Brustkrebs lässt sich an verschiedenen Parametern wie z. B. der lokalen und systemischen Therapie, aber insbesondere dem Langzeitoutcome der Patientinnen ermessen. Die Bundesgeschäftsstelle Qualitätssicherung gGmbH (BQS) [14] erfasst zwar Komplikationen und Qualitätsparameter, jedoch keine Follow-up-Informationen im Rahmen einer Brustkrebstherapie. Allerdings werden solche Parameter im Rahmen verschiedener Zertifizierungen (z. B. EUSOMA) [15] zwingend erforderlich sein. Dies führt zu der Situation, dass voraussichtlich viele Zentren retrospektive Kohortenanalysen ohne ein entsprechendes Datenmonitoring durchführen werden.



**Abb. 4** Simulationsexperiment zur Überprüfung des Einflusses eines möglichen Information-Bias. Kaplan-Meier-Analyse des Gesamtüberlebens der Patientinnen aufgrund der vorliegenden Follow-up-Informationen (Kurve A) und hypothetischer bester Verlauf (Kurve B), wenn früh zensierte Patientinnen kein Ereignis aufweisen.

Unsere Daten zeigen, dass die retrospektive Analyse des Überlebens der Patientinnen zahlreichen Confoundern unterliegt, was unabhängig von Qualitätsparametern eine Vergleichbarkeit mit anderen, großen Kollektiven erschwert und somit unter Umständen schlechtere Therapieergebnisse widerspiegelt. Die Erfassung des Krankheitsverlaufs mittels eines Erhebungsbogens erbrachte eine Antwortrate von 48,6%, eine Größenordnung, die auch in der Literatur beobachtet wurde [16]. Bereits bei der Beschreibung der klinischen Patientencharakteristika zeigte sich, dass in der Universitätsfrauenklinik vorzugsweise ein Kollektiv mit einem höheren Risiko für das Auftreten eines Krankheitsrezidivs in die Analyse Eingang gefunden hatte (größere Häufigkeit von jüngeren Patientinnen, größeren Tumoren und positivem Nodalstatus). Ob diese Häufung ein realer Effekt ist oder ein Bias aufgrund fehlender Informationen anderer Patientinnen entsteht, die aufgrund eines niedrigeren Risikoprofils im weiteren Follow-up nicht mehr an die Universitätsfrauenklinik angebunden waren oder keine Nachsorge haben durchführen lassen, muss dahingestellt bleiben. Als weiteres Ergebnis der Untersuchung war zu beobachten, dass sich beim Vergleich der Follow-up-Zeiten von Patienten mit und ohne Ereignis eine starke Häufung von Patientinnen zeigte, die sehr früh ohne Ereignis zensiert wurden. Dies wies im Vergleich zu anderen Populationskollektiven [11, 12] auf einen Information-Bias durch häufigeren „loss of follow-up“ in der Gruppe der krankheitsfreien Patientinnen hin. Das Ergebnis eines Simulationsexperiments mit dem Ziel, diesen Information-Bias abzuschätzen, lässt eine Größe von bis zu 5% Unterschied im 5-Jahres-Gesamtüberleben möglich erscheinen.



**Abb. 5a und b** Simulationsexperiment zur Überprüfung des Einflusses eines möglichen Information-Bias mit Stratifikation nach Prognosefaktor. Kaplan-Meier-Analyse des Gesamtüberlebens der Patientinnen aufgrund der vorliegenden Follow-up-Informationen (Abb. a) und hypothetischer bester Verlauf (Abb. b), wenn früh zensierte Patientinnen kein Ereignis aufweisen. Zusätzlich wurden die Patientinnen hier in a und b aufgrund des Tumorstadiums (T1–T4) stratifiziert.



Unsere Ergebnisse zeigen, welche kritischen Einflüsse bei retrospektiver Erfassung des Follow-ups von Mammakarzinompatientinnen zu beobachten und beim Vergleich solcher Daten ins Kalkül zu ziehen sind. Um eine wirkliche Beurteilung der Therapiequalität zu erhalten, sollten unbedingt prospektiv geplante Nachbeobachtungen angewendet werden. Auch wenn der Aufbau nationaler Krebsregister gut voranschreitet und das Mammakarzinom eine Krebserkrankung ist, die in den bevölkerungsbezogenen Krebsregistern am besten erfasst wird [17], so sind diese rein epidemiologischen Daten nicht ausreichend und ein flächendeckendes klinisches Krebsregister wäre eine logische Konsequenz. Zudem bestätigen unsere Ergebnisse die Wichtigkeit, bei Veröffentlichung von Überlebensdaten immer auch die detaillierten klinischen Charakteristika der Patientinnen einzuschließen und anzuführen, wie die Daten erhoben wurden bzw. wie die Antwortrate bei schriftlichen oder mündlichen Erhebungen war.

### Interessenkonflikt



Die Autoren erklären, dass sie keine finanziellen Verbindungen mit einer für den Artikel relevanten Firma haben.

### Literatur

- 1 Kreienberg R, Kopp I, Albert U et al. Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms. Deutsche Krebsgesellschaft e.V., Informationszentrum für Standards in der Onkologie (ISTO), Hrsg. Berlin: W. zuchschwerdt Verlag GmbH; 2008: ISBN 978-3-88603-934-0
- 2 Harnett A, Smallwood J, Titshall V et al., *Guideline Development Group*. Diagnosis and treatment of early breast cancer, including locally advanced disease – summary of NICE guidance. *BMJ* 2009; 338: b438
- 3 Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. *Lancet* 2002; 359: 341–345
- 4 Weiss C. Basiswissen Medizinische Statistik. Heidelberg: Springer Medizin Verlag; 2005: ISBN-13 978-3-540-71460-6
- 5 Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002; 359: 248–252
- 6 Mamdani M, Sykora K, Li P et al. Reader's guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. *BMJ* 2005; 330: 960–962
- 7 Mathis S, Gartlehner G. Kohortenstudien. *Wien Med Wochenschr* 2008; 158/5–6: 174–179
- 8 Lundin J, Lehtimäki T, Lundin M et al. Generalisability of survival estimates for patients with breast cancer – a comparison across two population-based series. *Eur J Cancer* 2006; 42: 3228–3235
- 9 Surveillance, Epidemiology, and End Results (SEER) Program Public-Use Data (1973–2001). National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program, Cancer Statistics Branch, released April based on the November 2003 submission (<http://www.seer.cancer.gov>)
- 10 Karn T, Metzler D, Ruckhäberle E et al. Data driven derivation of cutoffs from a pool of 3,030 Affymetrix arrays to stratify distinct clinical types of breast cancer. *Breast Cancer Res Treat* 2009; May 20 [Epub ahead of print]. DOI: 10.1007/s10549-009-0416-z
- 11 Miller LD, Smeds J, George J et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* 2005; 102: 13550–13555
- 12 Pawitan Y, Bjohle J, Amler L et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005; 7: R953–R964
- 13 Desmedt C, Piette F, Loi S et al., *TRANSBIG Consortium*. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* 2007; 13: 3207–3214
- 14 BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH. <http://www.bqs-online.com>
- 15 The European Society of Breast Cancer Specialists. <http://www.eusoma.org>
- 16 Nakash RA, Hutton JL, Jørstad-Stein EC et al. Maximising response to postal questionnaires – a systematic review of randomised trials in health research. *BMC Med Res Methodol* 2006; 6: 5
- 17 Robert Koch-Institut, *Gesellschaft der epidemiologischen Krebsregister*, Hrsg. Krebs in Deutschland 2003–2004. Häufigkeiten und Trends. Eine gemeinsame Veröffentlichung des Robert Koch-Instituts und der Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. 6. überarbeitete Auflage. Berlin: Druckhaus Berlin-Mitte GmbH; 2008: ISBN 978-3-89606-182-9